

Research Project Questions

DATA5207: Data analysis in the social sciences

Saaz Kaur Sahdra

Assessment overview

Seventy per cent of your final grade for this unit comes from the completion of the major research project.

This assessment is divided into two parts. *First*, you will be required to develop an 800-word research plan. This is due half-way through the unit (see canvas for due dates) and requires you to outline how you intend to complete the project. This is designed to allow you to gain early feedback on your approach to this assessment to increase your chances of passing as much as possible.

Second, there is a final paper due at the end of semester. This is a 2500-word research report. You must provide a more developed (but still brief) literature review, outline your data and methodology and specify why you have chosen the methods used, and present your results.

These assessments are designed to track your competence in the core skills developed in this unit, and provide you with the opportunity to apply them on practical research tasks.

For assessment purposes this item is considered your final exam. Failure to achieve at least a combined 40 per cent on the plan and project will result in a fail grade for the unit.

There are four options for topics to examine for this assessment. Once you have selected your question, you are expected to write and test your own theory answering the question, and examine the data using your chosen methods, providing you with flexibility in how you can approach the question (see the rubric, templates and guides on canvas for more information on what we expect).

All work for the plan and final report, including data cleaning and analysis, must be conducted in an *R* Markdown (RMD) file with your literature review and analysis (templates are provided for both assessments).

You can (and are encouraged) to look for additional data sources and other information, but this is not a requirement. Creativity and effort will be rewarded. Grades will be given for quality of analysis and presentation, and how well you use the methods and material covered in this class.

The research plan

For this first part of the assessment, you will need to outline your hypothesis and the working theory you plan to test, the academic literature that informs it (ten sources are required for your plan) and then outline the methodology you intend to use to answer it.

You should read through the canvas page to make sure you include what you believe will be relevant methods dealt with later in the semester, and discuss your ideas with the teaching team. No analysis will need to be conducted for this assessment. You just need to show you have thought about the question, the data you are using, and how you will examine it. **Although you should examine the data and familiarise yourself with the variables you are going to use, to make sure your approach makes sense and will work.**

Additionally, you are not locked in to this plan for your final research project. This provides an opportunity for you to think about and receive feedback on your research design before you complete your final assessment.

Questions

1. Was there a change in Australian attitudes towards climate before and after the 2019 fires? If so, what are predictors of shifted opinions?

The bushfire season in eastern Australia began early in 2019, with several serious fires across Northern Australia in June, followed by the unprecedented burning of subtropical rainforest in Queensland in September. In the following months the fire moved along the coast of New South Wales, and reached East Victoria by late December. Sydney and Canberra were shrouded in smoke for large parts of Spring and Summer. Breathing masks sold out in Melbourne in January, when it too was covered by ash from the nearby fires. Most experts say the hot and dry conditions that made these fires so severe were exacerbated by climate change. Others — mostly non-experts — blamed arson, or a lack of hazard reduction burns. This leads us to wonder if the scale and intensity of these fires, and the widespread impact from smoke across several of Australia's largest metropolitan areas, may have shifted public opinion on whether climate change was occurring in Australia. Using the data provided, you are to test whether attitudes towards climate change shifted between surveys, and what predicts changed attitudes.

Data provided: Survey data collected in July 2019 and January-February 2020 (folder contains .sav file and code book); just before and after the east coast fires. These are panel data with the same respondents interviewed for both surveys, which have been merged into a single file. You have two possible dependent variables that can be taken from these data. Questions Q29a/Q29_a from the 2019/20 surveys provide responses to questions about whether climate change is happening, and Q29b/Q29_b and Q29c/Q29_c if it is caused by human activity (whether a respondent received Q29 b or c depended on their answer to a). When you are trying to make sense of the survey, everything above line 6487 in the codebook file is from the 2019 survey, everything below this is from the 2020 re-contact, starting with the item `case_recon`.

2. What factors explain support for gun control in the United States?

Use the provided survey data, as well as other sources of information you can obtain, to establish a theory of support or opposition to gun control and an explanatory model.

Data provided: The General Social Surveys, 1972-2016, survey data file. More detail on this can be found [here](#). *Suggested additional data:* You can look at other survey data available (Pew Research has an excellent collection, much of which can be easily accessed). Additionally, you could also look at relationship between attitudes and rates of gun crime and other homicides, by state (or other geographical patterns). These data are available from a number of sources.

3. What are the predictors for better health outcomes at the county level in the United States?

Use a combination of the data provided and data from other sources to establish a theory of ecological health outcomes and an explanatory model. Your dependent variable will be the Premature age-adjusted mortality variable in sheet 6 of the provided dataset. You can use other variables from this dataset as predictors, as long as they make sense conceptually. You may want to (but are not required to) obtain data from other sources to help with your analysis.

Data provided: Data file of different health outcomes at the level of US counties, from the Robert Wood Johnson Foundation and the University of Wisconsin Population Health Institute. Some additional detail on this can be found [here](#). *Suggested additional data:* You may source additional predictors from the [Centers for Disease Control](#) and [the US Census Bureau](#).

4. What explains domestic violence rates in NSW local government areas?

Use a combination of the data sources provided, to establish a theory of ecological crime rates and an explanatory model. Your dependent variable are those cases listed as *Domestic violence related assault* in the variable Subcategory. Note: these data can be difficult to work with.

Data provided: Data file of reported incidents of domestic violence by local government area. This comes from the [Bureau of Crime Statistics and Research](#). *Suggested additional data:* You will need to source your own independent variables for this question (matching on LGA). You may want to examine other census data, and perhaps surveys that cover domestic violence. A good place to get data is the Australian Bureau of Statistics [Data Packs](#).